

# On the optimal control problem for a class of monotone bilinear systems

Neil K. Dhingra<sup>☆a,\*</sup>, Marcello Colombino<sup>☆☆b</sup>, Mihailo R. Jovanović<sup>☆a</sup>,  
Anders Rantzer<sup>c</sup>, and Roy S. Smith<sup>☆☆b</sup>

<sup>a</sup>*Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455*

<sup>b</sup>*Automatic Control Laboratory, Swiss Federal Institute of Technology, Zürich, Switzerland*

<sup>c</sup>*Automatic Control LTH, Lund University, SE-221 00 Lund, Sweden*

---

## Abstract

We consider a class of monotone systems in which the control signal multiplies the state. Among other applications, such bilinear systems can be used to model the evolutionary dynamics of HIV in the presence of combination drug therapy. For this class of systems, we formulate an infinite horizon optimal control problem, prove that the optimal control signal is constant over time, and show that it can be computed by solving a finite-dimensional non-smooth convex optimization problem. We provide an explicit expression for the subdifferential set of the objective function and use a subgradient algorithm to design the optimal controller. We further extend our results to characterize the optimal robust controller for systems with uncertain dynamics and show that computing the robust controller is no harder than computing the nominal controller. We illustrate our results with an example motivated by combination drug therapy.

*Keywords:* Positive systems, Decentralized control, Monotone systems, Optimal control, Structured control

---

## 1. Introduction

This work is motivated by recent developments on the modeling and design of combination drug therapy design for HIV treatment [1, 2, 3, 4, 5, 6, 7, 8, 9]. These advancements provide a framework for modeling the evolutionary dynamics of the population of different HIV mutants within an infected patient as a bilinear monotone system in which drug therapy affects the death rate of specific mutagens. The drug cocktail prescribed to the patient over the course of their treatment can be interpreted as a time-varying decentralized controller for the positive system that describes the virus mutation.

A system is called positive if, for every nonnegative initial condition and nonnegative input signal, its state and output remain nonnegative for all time [10]. Positive systems and their nonlinear counterpart, monotone systems, arise naturally when modeling physical and biological systems with inherently nonnegative states, e.g., the viral population of HIV. Other common examples include Markov chains and models of chemical

---

<sup>☆</sup>Supported in part by the National Science Foundation under Awards ECCS-1407958 and CNS-1544887.

<sup>☆☆</sup>Supported in part by the Swiss National Science Foundation under Award 2-773337-12.

\*Corresponding author

reaction networks, transportation networks, population dynamics, and heat transfer; in these, the states represent probabilities, concentrations, traffic density, species counts, and temperature, respectively.

Positive systems are not only interesting from a practical point of view but they also have rich structure and subtle system-theoretic properties. The study of positive systems dates back to the early 1900s, when Perron and Frobenius explored the spectral properties of nonnegative matrices. In recent years positive systems have gained renewed attention because several problems that are intractable in general greatly simplify for this class of systems.

It has been recently shown that the design of unconstrained decentralized controllers for positive systems is convex [11, 12]. However, many problems, such as the design of drug therapy for HIV, impose additional structural constraints which cannot be handled by the LMI approach proposed in [12]. To overcome this challenge, a design of  $\mathcal{L}_1$  and  $\mathcal{H}_\infty$  controllers which satisfy such structural constraints but achieve suboptimal performance has been examined in [5, 6, 7]. More recently, building on [3, 4], convexity of the structured decentralized  $\mathcal{H}_2$  and  $\mathcal{H}_\infty$  optimal control problems for positive systems was established in [9].

While references [5, 6, 7, 8, 13, 9] either assume a constant control signal or use heuristics to introduce time dependence, we show that such a constant input is in fact optimal for the induced power norm. We cast the optimal synthesis problem of constant control inputs as a finite-dimensional non-smooth convex optimization problem and develop an algorithm for designing the optimal controller. Finally, we exploit the advancements in [14, 15] to design robust controllers which guard against model uncertainty.

Our presentation is organized as follows. In Section 2, we formulate the optimal control problem for the class of bilinear positive systems that we study. In Section 3, we prove that a constant control input solves our optimal control problem and develop a subgradient algorithm. In Section 4, we show that designing the robust controller for systems with model uncertainty is no harder than computing a nominal controller. In Section 5, we apply our results to an example inspired by combination drug therapy design for HIV. Finally, we conclude the paper in Section 6.

## 2. Preliminaries and problem formulation

In this section, we provide necessary background material, introduce the class of bilinear positive systems that we study, and formulate the optimal control problem.

### 2.1. Preliminaries

The set of real numbers is denoted by  $\mathbb{R}$  and  $\mathbb{R}_+$  ( $\mathbb{R}_{++}$ ) is the set of nonnegative (positive) reals. The set of  $n \times n$  Metzler matrices (matrices with nonnegative off-diagonal elements) is denoted by  $\mathbb{M}^n$ . The set of  $n \times n$  nonnegative (positive) diagonal matrices is denoted by  $\mathbb{D}_+^n$  ( $\mathbb{D}_{++}^n$ ). We use  $\bar{\sigma}(A)$  to indicate the largest singular value of  $A$ .

The space of square integrable signals is denoted by  $\mathcal{L}_2$ . The inner product in this space is given by

$$\langle u, v \rangle_2 := \int_0^\infty u^T(t) v(t) dt$$

with the associated norm  $\|v\|_2^2 = \langle v, v \rangle_2$ . The power semi-norm of a signal  $v$  is

$$\|v\|_{\text{pow}}^2 := \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T v^T(t) v(t) dt. \quad (1)$$

The space of trigonometric polynomials is defined as

$$\mathcal{T} := \left\{ g : \mathbb{R} \rightarrow \mathbb{R}^n \left| g(t) = \sum_{k=1}^N \alpha_k e^{j\lambda_k t}, \lambda_k \in \mathbb{R}, \alpha_k \in \mathbb{C} \right. \right\},$$

where  $j$  is the imaginary unit. The closure of  $\mathcal{T}$  with respect to the metric  $\|f - g\|_{\text{pow}}$  is given by the space of Besicovitch almost periodic functions  $\mathcal{B}_2$  [16]. This is a Hilbert space with the inner product [17, 18]

$$\langle u, v \rangle := \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T u^T(t) v(t) dt$$

and the norm  $\|\cdot\|_{\text{pow}}$ . The mean of a signal  $v \in \mathcal{B}_2$

$$M(v) := \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T v(t) dt$$

is well defined for every  $v \in \mathcal{B}_2$ . Furthermore, each  $v \in \mathcal{B}_2$  can be decomposed uniquely as  $v = \bar{v} + \tilde{v}$ , where  $\bar{v}$  is a constant signal given by  $\bar{v} = M(v)$  and  $M(\tilde{v}) = 0$ . Note that the inner product between a constant signal  $\bar{v}$  and a zero-mean signal  $\tilde{v}$  is zero, i.e.,  $\langle \bar{v}, \tilde{v} \rangle = 0$ .

The space  $\mathcal{B}_2$  contains all bounded  $\mathcal{L}_2$  signals, periodic signals, and almost periodic signals. At the same time, it alleviates challenges arising from the fact that the space of signals with bounded power norm is not a Hilbert space; for additional discussion see [19].

## 2.2. Problem formulation

Consider the bilinear system

$$\dot{x} = (A + D(u))x + B d \quad (2a)$$

where  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times q}$ , and  $D: \mathbb{R}^m \rightarrow \mathbb{D}^n$  is a linear operator. For given control and disturbance signals  $u \in \mathcal{B}_2$  and  $d \in \mathcal{B}_2$ , we associate the performance output,

$$z_{u,d} = \begin{bmatrix} Q^{\frac{1}{2}} \\ 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ R^{\frac{1}{2}} \end{bmatrix} u \quad (2b)$$

with (2a), where  $Q \succeq 0$  and  $R \succ 0$  are the state and control weights.

**Assumption 1.** *A is a Metzler matrix,  $D: \mathbb{R}^m \rightarrow \mathbb{D}^n$  is a linear operator, B and Q are matrices with nonnegative entries, and there is a constant stabilizing control input  $\bar{u}$  for (2a).*

Under Assumption 1, (2) is a positive system. This implies that for every control input  $u$ , every nonnegative disturbance  $d$ , and every nonnegative initial condition  $x(0)$ , the state  $x$  and the output  $z_{u,d}$  of system (2) remain nonnegative at all times.

The induced power norm of a stable system (2) is,

$$J(u) := \sup_{\|d\|_{\text{pow}}^2 \leq 1} \|z_{u,d}\|_{\text{pow}}^2 \quad (3)$$

and it quantifies the response to the worst case persistent disturbance  $d$ . For unstable open-loop systems (2), there may be no stabilizing control input  $u$  in  $\mathcal{L}_2$ . Thus, the  $\mathcal{L}_2$ -induced gain does not provide a suitable measure of input-output amplification for (2) and  $J(u)$  represents an appropriate generalization of the  $\mathcal{H}_\infty$  norm for this class of bilinear positive systems.

We now formulate the optimal control problem.

**Problem 1.** *Design a stabilizing bounded control signal  $u \in \mathcal{B}_2$  to minimize  $J(u)$  for bilinear positive system (2).*

### 3. Solution to the optimal control problem

In this section, we prove that a constant control input solves Problem 1 and provide a subgradient algorithm for finding the optimal solution.

#### 3.1. Main result

Since  $J(u)$  is given by (3), any  $u^*$  which solves Problem 1 satisfies  $\|z_{u^*,d}\|_{\text{pow}}^2 \leq J(u^*) \leq J(u)$  for all  $u \in \mathcal{B}_2$  and  $d \in \mathcal{B}_2$ . In particular,  $\|z_{u^*,d}\|_{\text{pow}}^2 \leq J(\bar{u})$  where  $\bar{u}$  is a constant control input. As shown in [20], for constant control inputs the worst-case disturbance  $\bar{d}$  is also constant, i.e.,  $J(\bar{u}) = \|z_{\bar{u},\bar{d}}\|_{\text{pow}}^2$ . In what follows, we show that  $\|z_{u,\bar{d}}\|_{\text{pow}}^2$  is a convex function of  $u$ , that a constant  $\bar{u}$  minimizes it, and, thus, that a constant control input solves Problem 1.

We first establish convexity of  $\|z_{u,\bar{d}}\|_{\text{pow}}^2$ .

**Lemma 1.** *Let  $d(t) = \bar{d}$  be a constant non-negative disturbance. Then, under Assumption 1, the power norm of the output  $\|z_{u,\bar{d}}\|_{\text{pow}}^2$  is a convex function of  $u \in \mathcal{B}_2$ .*

PROOF. By [3, Theorem 2] every component of  $x(t)$  is a convex function of the control input  $u \in \mathcal{L}_2[0, t]$ . The power norm of the performance output (2b) is given by,

$$\|z_{u,\bar{d}}\|_{\text{pow}}^2 = \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T (x^T(t)Qx(t) + u^T(t)Ru(t)) dt.$$

The matrix  $R$  is positive definite so the second term on the right-hand side is a convex function of  $u$ . Furthermore, since the matrix  $Q \succeq 0$  has nonnegative entries,  $v^T Q v$  is nondecreasing in the elements of

$v$ . Thus, the composition rules for convex functions [21] imply that  $x^T(t)Qx(t)$  is a convex function of  $u \in \mathcal{L}_2[0, t]$ .

Let  $P_T$  denote the mapping that truncates the support of a signal  $v$  to  $[0, T]$ . If there is  $T$  such that  $P_T v$  is not an  $\mathcal{L}_2$  signal, the power seminorm (1) of  $v$  is infinite. This implies that, for any  $v \in \mathcal{B}_2$ ,  $P_T v \in \mathcal{L}_2[0, T]$ . Thus, any  $v \in \mathcal{B}_2$  can be written as  $\lim_{T \rightarrow \infty} P_T v$  of  $P_T v \in \mathcal{L}_2[0, T]$ . Since both the integral and the limit preserve convexity,  $\|z_{u, \bar{d}}\|_{\text{pow}}^2$  is a convex function of  $u \in \mathcal{B}_2$ .

**Lemma 2.** *Let  $\bar{u}$  be a stabilizing constant control input for (2a) and  $\bar{d}$  be a constant non-negative disturbance. Then, the directional derivative of  $\|z_{u, \bar{d}}\|_{\text{pow}}^2$  evaluated at  $\bar{u}$  is zero for any bounded zero-mean variation  $\tilde{u} \in \mathcal{B}_2$ .*

PROOF. The dynamics (2a) with control input  $\bar{u} + \varepsilon \tilde{u}$  and constant disturbance  $\bar{d}$  are

$$\dot{x} = (A + D(\bar{u} + \varepsilon \tilde{u}))x + B\bar{d}. \quad (4)$$

Since the unperturbed system (with  $\varepsilon = 0$ ) is exponentially stable and the solution  $x(t)$  is continuous in  $\varepsilon$ , (4) represents a system in a regularly perturbed form [22, 23]. The Taylor series expansion can be used to write the solution to the perturbed dynamics as,

$$x(t) = \bar{x}(t) + \varepsilon \tilde{x}(t) + \mathcal{O}(\varepsilon^2) \quad (5)$$

where  $\bar{x}$  is the nominal solution that solves (4) for  $\varepsilon = 0$

$$\dot{\bar{x}} = (A + D(\bar{u}))\bar{x} + B\bar{d} \quad (6a)$$

and  $\tilde{x}$  represents the first-order correction. By [23, Theorem 10.2],  $\tilde{x}$  is determined by the solution to a differential equation corresponding to the  $\mathcal{O}(\varepsilon)$  terms in the expression obtained by substituting (5) into equation (4),

$$\dot{\tilde{x}} = (A + D(\bar{u}))\tilde{x} + D(\tilde{u})\bar{x}. \quad (6b)$$

Both (6a) and (6b) are stable LTI systems. Thus, for the constant disturbance  $d(t) = \bar{d}$ , the solution  $\bar{x}$  to (6a) is asymptotically constant. Since  $\tilde{u}$  is zero-mean, the signal  $D(\tilde{u})\bar{x}$  is also zero-mean. The dynamics (6b) are a stable LTI system forced by a zero-mean input and therefore  $\tilde{x}$  is also zero-mean.

Because  $\bar{x}$  is asymptotically constant and  $\tilde{x}$  is zero-mean,  $\langle Q^{1/2}\bar{x}, Q^{1/2}\tilde{x} \rangle = 0$ . Similarly,  $\langle R^{1/2}\bar{u}, R^{1/2}\tilde{u} \rangle = 0$ . Furthermore,  $\|z_{u, \bar{d}}\|_{\text{pow}}^2 = \langle Q^{1/2}x, Q^{1/2}x \rangle + \langle R^{1/2}u, R^{1/2}u \rangle$ , and we have,

$$\begin{aligned} \|z_{(\bar{u} + \varepsilon \tilde{u}), \bar{d}}\|_{\text{pow}}^2 - \|z_{\bar{u}, \bar{d}}\|_{\text{pow}}^2 &= 2\varepsilon (\langle Q^{1/2}\bar{x}, Q^{1/2}\tilde{x} \rangle + \langle R^{1/2}\bar{u}, R^{1/2}\tilde{u} \rangle) + \mathcal{O}(\varepsilon^2) \\ &= \mathcal{O}(\varepsilon^2). \end{aligned} \quad (7)$$

Thus, the first order correction to  $\|z_{u, \bar{d}}\|_{\text{pow}}^2$  evaluated at a constant control input  $\bar{u}$  is zero, which completes

the proof.

**Remark 1.** *Lemma 2 does not need Assumption 1 and it holds for all bilinear systems of the form (2a) for which there is a stabilizing constant control input.*

Lemma 2 implies that a constant control signal  $u^*$

$$u^* \in \underset{\bar{u} \text{ constant}}{\operatorname{argmin}} \|z_{\bar{u}, \bar{d}}\|_{\text{pow}}^2,$$

provides a local minimum for  $\|z_{u, \bar{d}}\|_{\text{pow}}^2$ . Based on Lemma 1, this quantity is convex with respect to  $u$ . Thus,  $u^*$  is a global minimizer. The following theorem relates the power norm of the output of system (2) subject to a constant disturbance,  $\|z_{u, \bar{d}}\|_{\text{pow}}^2$ , with the worst case power norm amplification  $J(u)$  and shows that the constant control signal  $u^*$  solves Problem 1.

**Theorem 3.** *Let Assumption 1 hold. Then, a constant control input  $u(t) = u^*$  solves Problem 1.*

PROOF. Let  $u^*$  minimize  $J(u)$  over the space of *constant* functions. Since (2a) with a constant control input is an LTI system, the maximum power-amplification coincides with the  $\mathcal{H}_\infty$  norm. Moreover, because (2a) is a positive system, the maximal singular value of the frequency response matrix peaks at zero temporal frequency and the worst-case disturbance is a constant signal [20]. This implies that  $J(u^*) = \|z_{u^*, \bar{d}}\|_{\text{pow}}^2$  where  $\bar{d} = v \geq 0$  is a constant nonnegative disturbance and  $v$  is the right principal singular vector of the matrix  $-Q^{\frac{1}{2}}(A + D(u^*))^{-1}B$ .

Suppose there exists a time varying signal  $\hat{u} \in \mathcal{B}_2$  such that  $J(\hat{u}) < J(u^*)$ . Then, since  $J$  measures the worst-case disturbance amplification,  $\hat{u}$  must also decrease the power norm of the output of system (2) for a constant disturbance  $\bar{d} = v$ , i.e.,

$$\|z_{\hat{u}, \bar{d}}\|_{\text{pow}}^2 \leq J(\hat{u}) < \|z_{u^*, \bar{d}}\|_{\text{pow}}^2 = J(u^*). \quad (8)$$

We next show that this is not possible.

Every bounded  $u \in \mathcal{B}_2$  can be written as  $u = \bar{u} + \tilde{u}$  where  $\bar{u}$  is constant and  $\tilde{u}$  is bounded and zero-mean. Since  $u^*$  minimizes  $\|z_{u, \bar{d}}\|_{\text{pow}}^2$  over constant control inputs and Lemma 2 implies that  $u^*$  minimizes it over bounded zero-mean control inputs,  $u^*$  is a local minimizer of  $\|z_{u, \bar{d}}\|_{\text{pow}}^2$  over all bounded  $u \in \mathcal{B}_2$ . By Lemma 1, this implies that  $u^*$  is a global minimizer of  $\|z_{u, \bar{d}}\|_{\text{pow}}^2$ , contradicting (8) and completing the proof.

### 3.2. Subgradient algorithm

We next cast Problem 1 as a finite-dimensional convex optimization problem. For constant control inputs we derive a subdifferential set of the objective function and employ a subgradient method to compute an optimal solution. Since subgradient algorithms have slow convergence rate, we provide conditions under which the objective function is differentiable. In this case, a variety of standard tools including gradient descent and quasi-Newton methods can be readily employed to ensure faster convergence.

Theorem 3 implies that Problem 1 can be solved by minimizing  $J(u)$  over constant control inputs. As mentioned above, for a constant control input, (2a) is a positive system and the maximum power amplification is caused by a constant disturbance  $\bar{d} = v$  where  $v$  is the right principal singular vector of the matrix  $-Q^{\frac{1}{2}}(A + D(\bar{u}))^{-1}B$ . For a constant control input  $\bar{u}$ ,  $\|R^{\frac{1}{2}}u(t)\|_{\text{pow}}^2 = \bar{u}^T R \bar{u}$ , and the objective function is given by

$$J(\bar{u}) = \bar{\sigma}^2 \left( -Q^{\frac{1}{2}} (A + D(\bar{u}))^{-1} B \right) + \bar{u}^T R \bar{u}.$$

Since  $J(\bar{u})$  is not always differentiable, subgradient methods [24] can be used to find an optimal solution. Before we provide an explicit characterization for the subgradient set of  $J$ , we determine the adjoint of the linear operator  $D$ .

The adjoint of a linear operator  $D: \mathbb{R}^m \rightarrow \mathbb{D}^n \subset \mathbb{R}^{n \times n}$  is the linear operator  $D^\dagger: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^m$  which satisfies

$$\langle X, D(u) \rangle = \langle D^\dagger(X), u \rangle$$

for all  $u \in \mathbb{R}^m$  and  $X \in \mathbb{R}^{n \times n}$ . Without loss of generality, the operator  $D$  can be expressed as

$$D(u) = \sum_{k=1}^m D_k u_k = \text{diag}(D_u u)$$

where  $\text{diag}(x)$  denotes a diagonal matrix with the vector  $x$  on its main diagonal and the  $k$ th column of  $D_u \in \mathbb{R}^{n \times m}$  is given by the main diagonal of the matrix  $D_k \in \mathbb{D}^n$ . The adjoint of the operator  $D$  is then,

$$D^\dagger(X) = D_u^T \text{diag}(X) \quad (9)$$

where  $\text{diag}(X) \in \mathbb{R}^n$  is the vector determined by the diagonal elements of  $X$ .

**Proposition 4.** *Let  $D$  be a linear operator and  $A_{\text{cl}} := A + D(\bar{u})$  be Hurwitz. Then,*

$$\partial J(\bar{u}) = \left\{ 2\bar{\sigma}_{\text{cl}} \sum_i \alpha_i D^\dagger \left( A_{\text{cl}}^{-1} B v_i w_i^T Q^{1/2} A_{\text{cl}}^{-1} \right) + 2R\bar{u} \mid w_i^T \left( -Q^{1/2} A_{\text{cl}}^{-1} B \right) v_i = \bar{\sigma}_{\text{cl}}, \alpha \in \mathcal{P} \right\} \quad (10)$$

where  $\bar{\sigma}_{\text{cl}} := \bar{\sigma}(-Q^{1/2} A_{\text{cl}}^{-1} B)$  and

$$\mathcal{P} := \left\{ \alpha \mid \alpha_j \geq 0, \sum_j \alpha_j = 1 \right\}.$$

PROOF. The maximum singular value  $\bar{\sigma}(X)$  of a matrix can be expressed as [25],

$$\bar{\sigma}(X) = \sup_{\|w\|=1, \|v\|=1} w^T X v.$$

The subdifferential set of the supremum over a set of differentiable functions,

$$f(x) = \sup_{i \in \mathcal{I}} f_i(x)$$

is the convex hull of the subgradients of each function  $f_j$  that achieves the maximum [26, Theorem 1.13],

$$\partial f(x) = \sum_{j|f_j(x)=f(x)} \alpha_j \nabla f_j(x), \quad \alpha \in \mathcal{P}.$$

Therefore, the subgradient of  $\bar{\sigma}$  is given by

$$\partial \bar{\sigma}(X) = \left\{ \sum_j \alpha_j w_j v_j^T \mid w_j^T X v_j = \bar{\sigma}(X), \alpha \in \mathcal{P} \right\}.$$

The matricial derivative of  $X^{-1}$  and the application of the chain rule yield (10).

**Remark 2.** *Proposition 4 holds for any linear operator  $D: \mathbb{R}^m \rightarrow \mathbb{R}^{n \times n}$ . Since we are primarily concerned with operators  $D: \mathbb{R}^m \rightarrow \mathbb{D}^n$ ,  $D^\dagger$  in (10) can be replaced with (9).*

In general,  $J$  is nondifferentiable. To find a solution to Problem 1, we implement a subgradient algorithm [24],

$$u^{k+1} = u^k - \alpha_k d^k$$

where  $\alpha_k > 0$  is a step size and  $d^k \in \partial J(u^k)$  is a member of the subdifferential set of the objective function. For a small enough  $\alpha$ , this algorithm is guaranteed to converge with convergence rate  $\mathcal{O}(1/\sqrt{k})$ . Backtracking is used to ensure closed-loop stability, but, since the subgradient algorithm is not a descent method, sufficient descent criteria such as the Armijo rule may not be employed.

**Remark 3.** *When the graph associated with  $A$  is strongly connected,  $J$  is continuously differentiable [9, Proposition 9] and therefore a variety of methods, including gradient descent, Newton's method, and quasi-Newton methods [24] can be used to find the solution more efficiently.*

The computational complexity of each iteration in the subgradient algorithm is  $\mathcal{O}(n^3)$  because of matrix inversion and the singular value decomposition. This computational cost is of the the same order as that of evaluating  $J$ . Finally, we refer the interested reader to [9, Section VI], where we discuss the use of proximal algorithms [27, 28] to solve optimization problems of the form of Problem 1 augmented with structure-promoting nonsmooth regularizers.



#### 4. Robust optimal control problem

In this section we study the robust optimal control problem for system (2a) with model uncertainty. We consider uncertain dynamics of the form

$$\dot{x} = ((A + \Delta_A) + D(u - \delta_u))x + Bd, \quad (11)$$

where

$$\Delta_A = \begin{bmatrix} \delta_{11} & \cdots & \delta_{1n} \\ \vdots & \ddots & \vdots \\ \delta_{n1} & \cdots & \delta_{nn} \end{bmatrix}$$

represents the uncertainty in the matrix  $A$  and the vector  $\delta_u$  represents the uncertainty in the input  $u$ . The input uncertainty  $\delta_u$  is relevant in biological applications as explained in Section 5. We assume that the uncertainties are bounded as  $|\delta_{ij}| \leq \alpha_{ij}$  for all  $(i, j)$  and  $|\delta_{u_k}| \leq \beta_k$  with  $\alpha_{ij} \geq 0$  and  $\beta_k \geq 0$ . We define the set of admissible perturbations as

$$\mathbf{\Delta} := \{(\Delta_A, \delta_u) \mid |\delta_{ij}| \leq \alpha_{ij}, |\delta_{u_k}| \leq \beta_k\}$$

and

$$\tilde{A} := \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{bmatrix}, \quad \beta := \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}.$$

For a fixed  $\Delta_A$ ,  $\delta_u$  and  $u$ , we denote by  $J(u; \Delta_A, \delta_u)$  the power-induced gain of system (11). We are now ready to state the main result of this section.

**Theorem 5.** *Let Assumption 1 hold and let  $D: \mathbb{R}^m \rightarrow \mathbb{D}^n$  be such that  $-D(u) \in \mathbb{D}_+$  for all  $u \in \mathbb{R}_+$ . Provided there exists a  $\bar{u}$  such that  $(A + \tilde{A} - D(\beta)) + D(\bar{u})$  is Hurwitz, the solution to the robust optimal control problem*

$$\underset{u}{\text{minimize}} \quad \underset{(\Delta_A, \delta_u) \in \mathbf{\Delta}}{\text{maximize}} \quad J(u; \Delta_A, \delta_u)$$

*is given by the solution to Problem 1 applied to the system*

$$\dot{x} = ((A + \tilde{A} - D(\beta)) + D(\bar{u}))x + Bd.$$

We first state an established result which will be useful in proving Theorem 5.

**Lemma 6** ([29, Theorem 8.1]). *The following statements are true:*

a) *Let  $A \in \mathbb{R}^{n \times n}$  then  $|e^A| \leq e^{|A|} = |e^{|A|}|$ .*

b) Let  $A, B \in \mathbb{R}^{n \times n}$ , if  $0 \leq A \leq B$  then  $e^A \leq e^B$ .

c) Let  $A, B \in \mathbb{R}^{n \times n}$  if  $|A| \leq |B|$  then  $\bar{\sigma}(A) \leq \bar{\sigma}(B)$ .

PROOF (PROOF OF THEOREM 5). First we recall that we can write  $D(u) = \text{diag}(D_u u)$ . Given a constant  $u(t) = \bar{u}$ , define

$$\mathcal{A}_\Delta := (A + \Delta_A) + D(\bar{u} - \delta_u)$$

and

$$\mathcal{A} := (A + \tilde{A}) + D_u(\bar{u} - \beta).$$

Note that  $|\mathcal{A}_\Delta| \leq \mathcal{A}$ . We begin by showing that  $\mathcal{A}_\Delta$  is Hurwitz for all  $(\Delta_A, \delta_u) \in \mathbf{\Delta}$  if and only if  $\mathcal{A}$  is Hurwitz. Necessity is obvious as  $(\tilde{A}, \beta) \in \mathbf{\Delta}$ . To prove sufficiency, assume  $\mathcal{A}$  is Hurwitz. Since  $\mathcal{A}$  is Metzler, we know that there exists a vector  $p > 0$  such that

$$\mathcal{A}p < 0,$$

which is equivalent to

$$|a_{ii} + \alpha_{ii} + \bar{u}_i - \beta_i| p_i > \sum_{j \neq i} |a_{ij} + \alpha_{ij}| p_j.$$

Clearly for all  $(\Delta_A, \delta_u) \in \mathbf{\Delta}$ ,

$$|a_{ii} + \delta_{ii} + \bar{u}_i - \delta_{u_i}| p_i > \sum_{j \neq i} |a_{ij} + \delta_{ij}| p_j,$$

which means that  $\mathcal{A}_\Delta$  is *scaled diagonally dominant* and therefore Hurwitz.

Now that we know that  $\mathcal{A}_\Delta$  is stable for all  $(\Delta_A, \delta_u) \in \mathbf{\Delta}$ , we show that  $(\tilde{A}, \beta)$  is also the worst case perturbation for the performance index  $J$ . Let  $\alpha > 0$  be such that  $\mathcal{A} + \alpha I \geq 0$ . Note that  $|\mathcal{A}_\Delta + \alpha I| \leq \mathcal{A} + \alpha I$ . Let us consider  $J(u; \Delta_A, \delta_u)$  for a constant  $u(t) = \bar{u}$  such that  $\mathcal{A}$  is Hurwitz. Since  $\mathcal{A}_\Delta$  is also Hurwitz we conclude that

$$J(u; \Delta_A, \delta_u) = \bar{\sigma}^2 \left( Q^{\frac{1}{2}} \int_0^\infty e^{\mathcal{A}_\Delta t} dt B \right) + \bar{u}^T R \bar{u} \quad (12a)$$

$$= \bar{\sigma}^2 \left( Q^{\frac{1}{2}} \int_0^\infty e^{-\alpha t} e^{(\mathcal{A}_\Delta + \alpha I) t} dt B \right) + \bar{u}^T R \bar{u} \quad (12b)$$

$$\leq \bar{\sigma}^2 \left( Q^{\frac{1}{2}} \int_0^\infty e^{-\alpha t} e^{|\mathcal{A}_\Delta + \alpha I| t} dt B \right) + \bar{u}^T R \bar{u} \quad (12c)$$

$$\leq \bar{\sigma}^2 \left( Q^{\frac{1}{2}} \int_0^\infty e^{-\alpha t} e^{(\mathcal{A} + \alpha I) t} dt B \right) + \bar{u}^T R \bar{u} \quad (12d)$$

$$= J(u; \tilde{A}, \beta), \quad (12e)$$

where (12c) follows from the application of statements (a) and (b) in Lemma 6, while (12d) follows from the application of (b) and (c) of Lemma 6.

From (12) we know that, given any constant input  $u(t) = \bar{u}$ , the worst case perturbation  $(\Delta_A, \delta_u) \in \Delta$  is given by  $(\tilde{A}, \beta)$ . By Theorem 3 the optimal  $u(t)$  in response to the perturbation  $(\tilde{A}, \beta)$  is constant and therefore the proof is complete.

The statement of Theorem 5 is intuitive when considering a positive system. The worst case uncertainty is given when every component of  $\Delta_A$  hits its upper bound and every component of  $\delta_u$  hits its lower bound. Similar results without such intuitive solutions can be derived for more complex uncertainty descriptions where not all elements of  $\Delta_A$  and  $\delta_u$  are allowed to vary independently. For a deeper discussion on this topic we refer the reader to [15] or [30] where robust stability for positive systems is considered in a more general setting with coupled linear and nonlinear uncertainties, respectively.

## 5. Combination drug therapy for HIV

As shown in [1, 7], the mutation-replication dynamics of HIV can be modeled as bilinear system (2). Here, the  $i$ th component of the state vector  $x$  represents the population of the  $i$ th HIV mutant. The matrix  $A$  specifies the mutation and replication rates of the virus;  $A_{ij}$  represents the rate at which mutant  $i$  turns into mutant  $j$  and  $A_{ii}$  represents the net replication rate of mutant  $i$ . The control input  $u_k$  is the dose of drug  $k$  and  $D(e_k)$  specify how efficiently drug  $k$  neutralizes each HIV mutant. The exogenous input  $d$  represents the effect of noise, disturbances or unmodeled dynamics on the virus population. The performance output (2b) captures the lethality each virus strand via  $Q$  and the penalizes the magnitude of the drug dosages via  $R$ .

### 5.1. Uncertain model of HIV dynamics

A large challenge with HIV virus dynamics is model uncertainty. Uncertainty arises from two different aspects:

- It is difficult to accurately estimate replication and mutation rates. Moreover, these rates vary between patients.
- It is difficult to precisely deliver drugs. Maintaining the target drug concentration at the treatment site is itself a challenging control problem.

In dynamics (11),  $\Delta_A$  represents the uncertainty in the replication-mutation rates and  $\delta_u$  represents uncertainty in the delivered drug dosage.

Theorem 5 has a clear physical interpretation for this application; the worst case perturbation is the one for which the virus replicates and mutates most aggressively and the drugs are the least effective. Furthermore, Theorem 3 implies that the optimal treatment strategy for minimizing the effect of disturbances is maintaining a constant concentration of drugs.

### 5.2. An example

In this section we present a simple example to illustrate the utility of the robust optimal control formulation. Consider the system described in Section 5 for a virus with  $n$  mutagens whose mutation pattern is given by the path graph in Figure 1 where  $r$  represents the replication rate of each mutagen. The entry  $c \geq 0$  represents a mutation pathway which is nominally zero but whose value is uncertain, i.e.  $\Delta_A = c e_1 e_n^T$ . The control authority  $u$  represents a drug which treats all the mutagens identically with no uncertainty, i.e.  $D(u) = -uI$  and  $\beta = 0$ . The input disturbance matrix is given by  $B = e_n$ , the state penalty matrix is  $Q = e_n e_n^T$ , and the control input penalty matrix is  $R = \rho I$ .

Uncertainty in the model is natural when the mutation pathways are not completely understood and can be a convenient way to account for nonlinear dynamics. The uncertain mutation pathway from mutagen  $x_n$  to  $x_1$  is particularly important because a nonzero value introduces a cycle into the graph associated with  $A$ . The optimal drug dose designed for the nominal model may not be stabilizing when applied to the true dynamics.

Since  $A + D(u)$  is lower triangular when  $c = 0$ , its eigenvalues are the diagonal entries and therefore the closed-loop system is nominally stable when  $r - u < 0$ . When  $c \neq 0$ , the graph associated with  $A + D(u)$  is a strongly connected graph and therefore has a unique largest eigenvalue [9, Proposition 9]. It follows that any value of  $c$  which changes the sign of the determinant,

$$\det(A + D(u)) = (r - u)^n + (-1)^{n-1}c$$

relative to  $\det(A + D(u))$  when  $c = 0$ , i.e., any  $c$  such that

$$c \geq |(r - u)^n|$$

causes the system to be unstable.

By computing the  $n$ th entry of  $(A + D(u))^{-1}$ , the objective function in Problem 1 can be explicitly expressed as,

$$J(u) = \left( \frac{|(u - r)^{n-1}|}{|(u - r)^n - c|} \right)^2 + \rho u^2.$$

In the nominal case when  $c = 0$ ,  $J(u) = (r + u)^{-2} + \rho u^2$ , so the optimal value of  $u$  solves the quartic equation,

$$(r + u)^{-3} - \rho u = 0$$

We consider the system in Fig. 1 with  $n = 10$  mutants, replication rate  $r = 1$ , and control penalty  $\rho = 3$ . A nominally optimal drug dose  $u_{\text{nom}}$  was designed to minimize  $J(u)$  for the nominal model (i.e.,  $c = 0$ ) and a robust optimal drug dose  $u_{\text{rob}}$  was designed to solve the robust optimal control problem considered in Theorem 5 with the assumption  $c \leq 0.1$  (i.e.,  $\alpha_{1n} = 0.1$  and all other  $\alpha_{ij} = 0$ , or, equivalently,  $\tilde{A} = 0.1 e_1 e_n^T$ ).

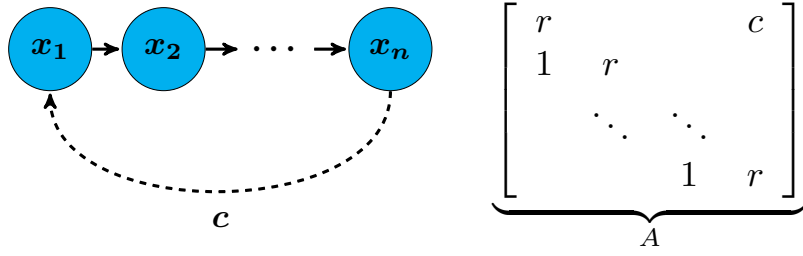


Figure 1: The dynamics are described by a path graph. The dashed line in the diagram and the  $c$  entry in the matrix represent an uncertain link from  $x_n$  to  $x_1$  which would introduce a cycle into the mutation pattern.

The designed doses were,

$$u_{\text{nom}} = 1.5936, \quad u_{\text{rob}} = 1.9413.$$

Any  $c \geq 0.0054$  destabilizes a system controlled by  $u_{\text{nom}}$  and any  $c \geq 0.5461$  destabilizes a system controlled by  $u_{\text{rob}}$ .

Figure 2a shows the impulse response for the nominal model controlled by  $u_{\text{nom}}$  and  $u_{\text{rob}}$ . Figure 2b shows the closed loop impulse response when  $c = 0.02$ . The nominal controller is more sensitive to perturbations in the model.

## 6. Concluding remarks

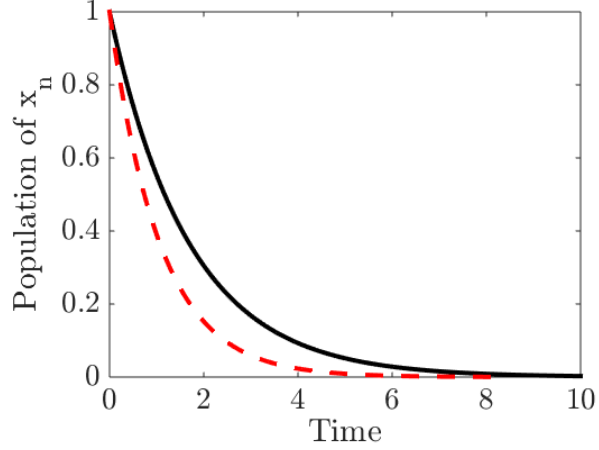
We study an infinite horizon optimal control problem for a class of monotone bilinear systems. The objective function is given by the induced power norm which represents a generalization of the  $\mathcal{H}_\infty$  performance metric. We prove that the optimal control signal is constant over time, show that the optimal controller can be computed by solving a finite-dimensional non-smooth convex optimization problem, and develop a subgradient algorithm to compute the optimal solution. We further extend our results for systems with uncertain dynamics and show that computing the optimal robust controller is no harder than computing the nominal controller. Finally, we apply our results to an example inspired by combination drug therapy design for the treatment of HIV.

## Acknowledgments

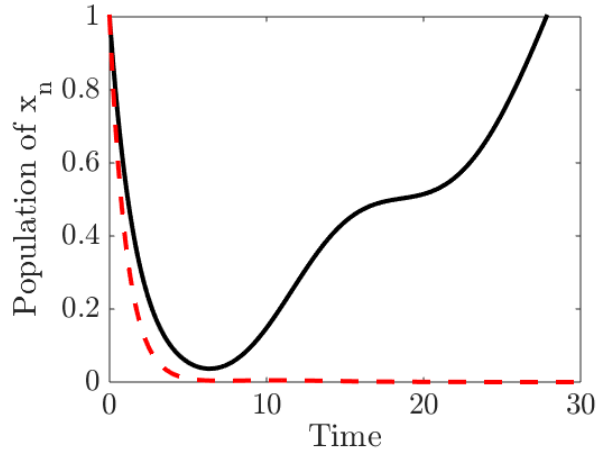
The authors would like to acknowledge Peyman Mohajerin Esfahani and Kevin Nowland for helpful discussions on functional analysis.

## References

- [1] E. Hernandez-Vargas, P. Colaneri, R. Middleton, F. Blanchini, Discrete-time control for switched positive systems with application to mitigating viral escape, *Int. J. Robust Nonlin.* 21 (10) (2011) 1093–1111.
- [2] E. A. Hernandez-Vargas, P. Colaneri, R. H. Middleton, Optimal therapy scheduling for a simplified HIV infection model, *Automatica* 49 (9) (2013) 2874–2880.



(a) Closed loop impulse responses of the nominal system



(b) Closed loop impulse responses of the system with  $c = 0.02$

Figure 2: Closed loop impulse response in the (a) nominal and (b) perturbed model with the nominal controller  $u_{\text{nom}}$  (**black** —) and the robust controller  $u_{\text{rob}}$  (**red** - -).

- [3] P. Colaneri, R. H. Middleton, Z. Chen, D. Caporale, F. Blanchini, Convexity of the cost functional in an optimal control problem for a class of positive switched systems, *Automatica* 4 (50) (2014) 1227–1234.
- [4] A. Rantzer, B. Bernhardsson, Control of convex monotone systems, in: *Proceedings of the 53rd IEEE Conference on Decision and Control*, 2014, pp. 2378–2383.
- [5] V. Jonsson, A. Rantzer, R. M. Murray, A scalable formulation for engineering combination therapies for evolutionary dynamics of disease, in: *Proceedings of the 2014 American Control Conference*, 2014, pp. 2771–2778.
- [6] V. Jonsson, N. Matni, R. M. Murray, Reverse engineering combination therapies for evolutionary dynamics of disease: An  $H_\infty$  approach, in: *Proceedings of the 52nd IEEE Conference on Decision and Control*, 2013, pp. 2060–2065.
- [7] V. Jonsson, N. Matni, R. M. Murray, Synthesizing combination therapies for evolutionary dynamics of disease for nonlinear pharmacodynamics, in: *Proceedings of the 53rd Conference on Decision and Control*, IEEE, 2014, pp. 2352–2358.

- [8] N. K. Dhingra, M. R. Jovanović, Convex synthesis of symmetric modifications to linear systems, in: Proceedings of the 2015 American Control Conference, Chicago, IL, 2015, pp. 3583–3588.
- [9] N. K. Dhingra, M. Colombino, M. R. Jovanović, On the convexity of a class of structured optimal control problems for positive systems, in: Proceedings of the 2016 European Control Conference, Aalborg, Denmark, 2016, pp. 825–830.
- [10] L. Farina, S. Rinaldi, Positive Linear Systems: Theory and Applications, John Wiley & Sons, 2011.
- [11] A. Rantzer, Scalable control of positive systems, Eur. J. Control 24 (2015) 72–80.
- [12] A. Rantzer, On the Kalman-Yakubovich-Popov lemma for positive systems, IEEE Trans. Automat. Control 61 (5) (2016) 1346–1349.
- [13] N. K. Dhingra, M. Colombino, M. R. Jovanović, Leader selection in directed networks, in: Proceedings of the 55th IEEE Conference on Decision and Control, Las Vegas, NV, 2016, to appear.
- [14] M. Colombino, R. Smith, Convex characterization of robust stability analysis and control synthesis for positive linear systems, in: Proceedings of the 53rd IEEE Conference on Decision and Control, Los Angeles, California, 2014, pp. 4379–4384.
- [15] M. Colombino, R. Smith, A convex characterization of robust stability for positive and positively dominated linear systems, IEEE Trans. Automat. Control 61 (7) (2016) 1965–1971.
- [16] A. S. Besicovitch, Almost Periodic Functions, Dover Publications, 1954.
- [17] C. Corduneanu, Almost periodic oscillations and waves, Springer Science & Business Media, 2009.
- [18] C. Constanda, M. E. Perez, Integral methods in science and engineering, Springer, 2010.
- [19] J. Mari, A counterexample in power signals space, IEEE Trans. Automat. Control 41 (1) (1996) 115–116.
- [20] T. Tanaka, C. Langbort, The bounded real lemma for internally positive systems and H-infinity structured static state feedback, IEEE Trans. Automat. Control 56 (9) (2011) 2218–2223.
- [21] S. Boyd, L. Vandenberghe, Convex optimization, Cambridge university press, 2004.
- [22] P. Kokotović, H. K. Khalil, J. O'Reilly, Singular perturbation methods in control: analysis and design, SIAM, 1999.
- [23] H. K. Khalil, Nonlinear Systems, 3rd Edition, Prentice Hall, 1996.
- [24] D. P. Bertsekas, Nonlinear programming, Athena Scientific, 1999.
- [25] S. Boyd, L. Vandenberghe, Convex optimization, Cambridge University Press, 2004.
- [26] N. Z. Shor, Minimization methods for non-differentiable functions, Vol. 3, Springer Science & Business Media, 2012.
- [27] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imaging Sci. 2 (1) (2009) 183–202.
- [28] N. Parikh, S. Boyd, Proximal algorithms, Foundations and Trends in Optimization 1 (3) (2013) 123–231.
- [29] R. A. Horn, C. R. Johnson, Matrix analysis, Cambridge university press, 1985.
- [30] M. Colombino, A. B. Hempel, R. S. Smith, Robust stability of a class of interconnected nonlinear positive systems, in: Proceedings of the 2015 American Control Conference, Chicago, IL, 2015, pp. 5312–5317.